Abstract

Hate speech on social media is a real problem with real consequences. Despite the constant efforts of social media platforms to moderate, flag, and ban hate posts, there is still a vast amount of hateful content flooding them. Hate speech, in general, and offensive material online, in particular, are not easy to define and may include a wide spectrum of expression. To thoroughly account for the nature and intensity of the effects of hate speech in social media requires to distinguish between various shades of hate speech targeting different groups and their subsequent effects. This paper seeks to review the literature on the psychological effects of online hate speech on the LGBT community and to highlight the strong negative impact of this phenomenon. The paper aims to contribute to the field by examining the propagation and the effects of derogatory language and hate speech based on sexual orientation.

Keywords: hate speech, social media, LGBT, derogatory speech

1. INTRODUCTION

Hate speech on social networking sites (SNS) is a widespread phenomenon affecting broad categories of users with the help of the low-cost and high-speed dissemination mechanism provided by these sites. Considerable work has been undertaken to combat harmful and abusive content, mainly through reactive measures. Hate speech is a term used to refer to any kind of offensive material, i.e., verbal, nonverbal, symbolic, or communicative actions that are deliberately used to denigrate and belittle members of a particular social group based on their membership (Simpson, 2013). Online hate speech expresses hatred of a collective that is different in terms of race, ethnicity, gender, religion, sexual orientation, political views, or other traits (Hawdon, Oksanen & Räsänen, 2017; Blazak, 2009). Unlike cyberbullying, where malice and degrading attitudes target individual, hate speech is addressed towards a collec-
tive to vilify it through verbal statements, nonverbal messages, symbols, images, or memes (Simpson, 2013).

There are several kinds of hate messages, including discriminatory speech, hate speech, incitement to hatred, incitement to terrorism, and incitement to genocide, listed by the intensity of the feeling (Ghanea, 2012). Internet intermediaries use disparate definitions of hate speech and guidelines to regulate it, employing different terms such as unlawful, harmful, threatening, abusive, harassing, tortuous, defamatory, vulgar, obscene, invasive, hateful, offensive, inappropriate speech that demeans or attacks a group based on race or ethnic origin, religion, disability, severe disease, gender, age, veteran status, and sexual orientation or gender identity. Although social networking sites make efforts to remove hate speech, they allow humor, satire, or social comments related to these topics (Gagliardone, 2015). Findings from a study conducted in the United Kingdom, United States, Finland, and Germany indicate that approximately 42% of 15- to 30-year-olds are exposed to hateful material online, especially on SNS (Keipi et al., 2016). Hate is not the only motivation that fuels hate speech. The sense of fear or loss, alienation, boredom, attention-seeking, pleasure in being controversial, economic self-interest are factors that can determine people to publicly insult or de-ride members of certain groups (Brown, 2017).

Previous research has shown that hate speech can have a detrimental impact on the targeted individuals or groups, ranging from causing psychological harm (Nielsen, 2002; Maitra, 2012; Boeckmann & Liew, 2002) to inciting violence (Muller & Schwarz, 2018; Fyfe, 2017). Hate speech offends, assaults the dignity, invades the autonomy, creates emotional distress, affects the personal development of the victims, and generally, by disrespecting the cultural diversity, undermines the democratic process (Brown, 2015). Hate speech exposure can have short-term consequences as mood swings, anger, loneliness, and fear, while long-term exposure can erode social trust, lead to radicalization and the perpetuation of extremist ideologies (Lee & Leets, 2002; Tynes, 2006; Foxman & Wolf, 2013; Hawdon, Oksanen & Räsänen, 2017). Although social networking sites by themselves do not increase exposure to online hate (Costello et al., 2018; Hawdon, Oksanen & Räsänen, 2014), they nonetheless provide the opportunity for people to consume hate speech online accidentally or deliberately, which increases the risk that they become ensnared in echo-chambers populated with like-minded people who share and reify their beliefs and their worldview (Costello et al., 2018).

Hate speech is rooted in ideologies of racism, sexism, religious intolerance, xenophobia, and homophobia, leading to victimization (Alkiviadou, 2018). Hate victimization starts with perceptions that are modified into stereotypes, leading then to prejudice and feelings of inferiority, exclusion, marginalization, de-prioritization, and, lastly, to victimization (Judge & Nel, 2018). A central part of the harm of hate speech is that it assaults one’s self-respect (Seglow, 2016). Moreover, hate speech and hate action go hand in hand, emphasizing how the discursive and the material dimensions of homophobia lead to one another (Judge, 2017). For example, empirical research on LGBT people indicates that hate speech and physical harm occur together in homophobic attacks (Nel & Judge, 2008; Breen & Nel, 2011).

Hate group websites usually include fantasy themes to maintain the hate, as well as abusive language towards the hated categories of people (Duffy, 2003; Hirvonen, 2013). Some common themes that are dramatized within hate groups refer to fairness, justice, morality used to justify hate claims and beliefs, the fantasy of a new and better world, the idea of the few chosen people that have to fight for fairness (Duffy, 2003). On the other hand, the hated subjects are depicted in offensive terms, and usually, a monstrous portrait is being created in
order to justify the need for action. Even the children are vilified to justify the hate towards the group (Hirvonen, 2013). The language is usually highly abusive and incites to violence. The lenses that are used to describe not only the children, but all the groups exposed to hate are those of criminality, mass threat, and dishonesty as dangerous secret plans are to be unfolded. Far-right websites promote a danger theme, which amplifies the violence or a fear theme, which stresses the statements that can incite the fear that the other group is planning to use violence (Gagliardone, 2015).

The choice of sexual orientation hate speech is motivated by the intense debate generated online (especially on Facebook) by the national referendum (held on October 7-8, 2018) to modify the Romanian Constitution to legally recognize as a family only the union between a man and a woman. The referendum did not pass, but it succeeded to polarize Romanians’ opinions regarding sexual minorities. Empirical research indicates that LGBT people are subjected to high levels of verbal assault that lead to considerable psychological harm (Walters et al., 2017; Breen & Nel, 2011; Nel & Judge, 2008; Polders et al., 2008). Trans people are even more likely to experience heightened levels of threat and vulnerability and face the lack of support of close ones compared with non-trans LGBT people (Walters et al., 2017).

This paper seeks to explore from a theoretical perspective the psychological effects of online hate speech on the LGBT community to highlight the strong negative impact of this phenomenon. Because a series of recent events in Romania brought to the public’s attention the problem of online abuse and discrimination of people from the LGBT community, the present paper aims to be a starting point for future extensive research regarding the challenge of LGBT online hate speech in Romania.

2. HATE SPEECH ON SOCIAL MEDIA

2.1. Emerging media – a perpetrator of hate?

Social media, through their low-cost, high-speed dissemination mechanism, facilitate the spread of hate speech. Brown (2017) makes a comparative analysis of online and offline hate speech and states that online hate speech could have worse effects because it takes place in front of a much larger audience, and victims may feel more ashamed. Social media offers the possibility to post hate comments anonymously. Compared to the offline environment, users can express messages of loathing without taking responsibility for it, and this encourages them to spread more contempt than they do face to face (Brown, 2018). It is proven that there is a positive correlation between anonymity and hate speech. Hate speech on race and sexual orientation is posted anonymously to a greater extent, compared to other categories of hate speech (Mondal, Silva & Benevenuto, 2017). Social media facilitates the creation and expansion of communities of like-minded people. This aspect also applies to hate speech because the Internet has a democratic character and offers the opportunity for every person to get involved in hate speech (Brown, 2018).

Online routine activities can increase the chances of victimization by bringing potential targets into contact with offenders in environments with low control or guardianship (Eck & Clarke, 2003; Reyns et al., 2011). Using social networking, file-sharing programs, instant messaging, sharing information through chat groups can increase the likelihood of harassment (Navarro & Jasinski, 2012). Those who expose personal information, add strangers
to their social networks, and have large numbers of friends or those who chat with strangers and confide in others anonymously become more vulnerable to victimization (Hawdon et al., 2015). Those who take part in online group deviance or who overexpose themselves through sexting are more likely to become cyberviolence targets (Costello et al., 2016). The connection with deviant peers associates with the probability of unwanted contact, sexual advances or harassment, and cyberstalking (Reyns et al., 2015).

The strong negative effects of online hate speech have led to the creation of mechanisms to reduce this type of abuse, especially through the use of algorithms. Social media companies have developed tools by which hate speech and users’ accounts who post hateful messages are blocked and deleted (Mathew et al., 2019). The expansion of the hate speech phenomenon in the online environment has also required the development of combat measures at the European level. Thus, through an agreement between the European Commission and Facebook, Microsoft, Twitter, and YouTube, it was created The EU Code of conduct on countering illegal hate speech online. Based on this hate code of conduct, the reported hate messages are deleted (The EU Code of conduct on countering illegal hate speech online, 2020). A very high percentage of messages inciting to hate are removed in this way, which shows that the European Commission’s initiative is beneficial.

Citron and Norton (2011) state that there are four ways a person can respond to hate speech: inaction, deletion, developing initiatives to educate social media users, or counter-speech. As Hubbard (2020) points out, hate speech victims tend to normalize abuse and do not report it. This aspect leads to the perpetuation of online hate speech. Also, the debates on combating hate speech (Mathew et al., 2019) emphasize that considering the punishment of spreading hate speech in the online environment is a violation of the right to free speech. Therefore, virtually anyone with access to the Internet could potentially produce, publish, and disseminate hateful materials affecting many people almost instantly under free speech rights.

2.2. Psychological effects of online hate speech

Hate speech can have substantial negative effects on victims’ mental health. For example, Tynes et al. (2008) show that online racial discrimination is strongly correlated with anxiety and depression. Brown (2015) emphasizes that hate speech has both short-term and long-term effects on victims’ mental health. As a short-term effect, the consequences of hate speech could translate into severe emotional distress manifested through anxiety, panic, shame, or fear. Hate speech can have similar effects to hate crimes, such as psychological trauma or negative effects on the community (Gerstenfeld, 2017). The long-term effects of hate speech include stress, psychosomatic disorders, anxiety, depression, alcoholism (Brown, 2015). Another severe aspect consists of pathological coping mechanisms that hate speech victims develop to cope with the abuse. The emotional suppression and the emotional transference are examples of inadequate coping strategies. Cognitive dissonance is another common phenomenon (Brown, 2015). To avoid severe emotional distress, hate speech victims develop maladaptive coping mechanisms to ignore the abuse they experience. However, this strategy leads to harmful effects on the well-being of victims (Brown, 2015).

In a study on anti-Muslim hate speech, Awan and Zempi (2016) demonstrate that victims experience a heightened state of fear. Victims of this type of abuse are always worried because they believe that threats received online could materialize offline at any time. Also,
since threats are launched anonymously, the victims fear that anyone could assault them, and this accentuates their desire to withdraw from society.

All in all, the harmful effects of online hate speech hateful are diverse. Other than being a precursor to possible hate crimes and violence, hate speech exposure can have psychological influences on morale, including heightened stress, anxiety, depression, and desensitization.

2.3. The case of the LGBT community

Despite various social networking sites’ attempts to eliminate hate speech, the LGBT community remains a target for users who spread hatred in the online environment. Prior to analyzing the psychological effects of hate speech on the LGBT community, it is necessary to point out that LGBT members have an increased risk of developing mental health problems. Previous research has identified that compared to heterosexuals, LGBT members have an increased risk of developing poor mental health, issues regarding excessive alcohol consumption, and smoking (Friedriksen-Goldsen et al., 2013). Also, transgender people are more likely to develop symptoms of depression and experience stress compared to non-transgender people in the LGBT community (Friedriksen-Goldsen et al., 2014). Transgender people are also more likely to experience hate crimes (Walters et al., 2017).

LGBT activists are targets of online hate speech, and this type of abuse has multiple psychological effects. In a study that focuses on the harmful effects of hate speech on LGBT activists in Moldova and Ukraine, the interviewees state that the online hate content causes them a series of harmful effects. These include emotional distress, depression, sleep disturbances, exhaustion, panic attacks that affect their daily activities, fear, or desire for social isolation (Nyman & Provozin, 2019). Hate messages strongly influence the decision of victims to delete or close their accounts on social networks as a defense mechanism. According to LGBT activists in Ukraine and Moldova, indirect hate speech has a similar negative effect to the direct one (Nyman & Provozin, 2019).

A recent study on online abuse experienced by members of the LGBT community highlights the impact that insults and threats have on mental health (Hubbard, 2020). In the case of the LGBT community, online violence causes a multitude of undesirable psychological effects from feelings of anger, sadness, anxiety, depression, stress, shame, self-blame, and social isolation. Victims of online abuse tend to feel guilty for being part of the LGBT community, and they isolate themselves to reduce the occurrence of hate speech. Online maltreatment also influences suicidal thoughts (Hubbard, 2020).

Dragowski et al. (2011) indicate that there is a strong correlation between experiencing verbal and physical abuse regarding sexual orientation and the manifestation of posttraumatic stress symptoms. Posttraumatic stress symptoms can include intensified avoidance behaviors. Abused people tend to avoid certain places, people, and situations that can lead to the appearance of negative thoughts and feelings, or reactive mechanisms (Parekh, 2017).

Because LGBT individuals are more prone to receiving hateful messages constantly than straight individuals, they end up internalizing negative attitudes toward the LGBT community (Ghafoori et al., 2019). The perpetual stigmatization of the LGBT community can lead to an internalized homophobia. Depression and PTSD are severe outcomes of internalized homophobia (Wotten, 2018). The internalization of these attitudes leads to the development of defense mechanisms such as hiding the gender identity or sexual orientation (Global Perspectives on the Trauma of Hate-Based Violence, 2020). The prevalence of online hate speech about
the LGBT community influences victims to develop self-censorship mechanisms and to suppress their gender identity or sexual orientation to be socially accepted (Berecz & Devinat, 2017). Thus, people who choose to hide their real sexual orientation or gender identity are forced to create a persona and live a double life, therefore registering a substantial negative impact on their mental health (Berecz & Devinat, 2017). It is essential to point out that transgender people have difficulties in this regard because the physical changes during the transition period are difficult to hide. This aspect could represent an explanation for the fact that transgender people feel more threatened, vulnerable, and anxious when considering the hate speech phenomenon (Walters et al., 2017).

Polders and his collaborators (2008) identified several factors that make members of the LGBT community more vulnerable to depression. However, their research shows that low self-esteem and frequent exposure to hate speech are the only factors strongly correlated with vulnerability to depression in the LGBT community. The correlation between hate speech and vulnerability to depression should come to the attention of the general public, especially since the level of suicide in the LGBT community is much higher compared to the rest of the population (Berecz & Devinat, 2017). Hate speech has destructive psychological effects on victims, and its prevalence in the online environment can also lead to desensitization (Soral, Bilewicz & Winiewski, 2017), thus decreasing the rate of reported and banned cruelties. At the same time, hate speech victims may begin to normalize the abuse and resort to avoidance as a coping mechanism (Hubbard, 2020).

Desensitization and normalization should constitute the subject of extensive research as they represent mechanisms that perpetuate the abuse on social media. It is not easy to detect and reduce hate speech by either employing algorithms or people, as it is highly sensitive to contextual and subjective judgments. A more proactive approach is needed to counter online hate speech based on a collective intelligence solution to identify and combat harmful content.

3. CONCLUSIONS

Online hate speech is a phenomenon that leads to harmful psychological effects, and its prevalence on social media is heightened. As the effects of hate speech range from anger to severe depression and suicide, studying the adverse effects and the factors that predict and moderate the consequences should be a priority.

As transgender people are most prone to experiencing online hate speech, and the effects more destructive in their case, this category should represent a central concern. Therefore, new research initiatives should address the particular way in which transgender people react and cope with hate speech online along with the countering measures. The desensitization to hate speech and the tendency to normalize the abuse seem to perpetuate the violence on social media, so future measures can include raising awareness on the topic and emphasizing new healthier coping mechanisms. Social media algorithms training could be used to identify better, classify, and censor online homophobic content. Machine learning should address the issue of user comments and satire as well.
Note

This paper was financially supported by the Human Capital Operational Program 2014-2020, co-financed by the European Social Fund, under the project POCU/380/6/13/124708 no. 37141/23.05.2019, with the title “Researcher-Entrepreneur on Labour Market in the Fields of Intelligent Specialization (CERT-ANTREP)”, coordinated by the National University of Political Studies and Public Administration.

References

Berecz, T., & Devinat, C. (2017). Relevance of Cyber Hate in Europe and Current Topics that Shape Online Hate Speech.
Brown, A. (2018). What is so special about online (as compared to offline) hate speech?. Ethnicities, 18(3), 297-326.
Hubbard, L. (2020) Online Hate Crime Report: Challenging online homophobia, biphobia and transphobia. London: Galop, the LGBT+ anti-violence charity


